

1

2

3

4

5

6

7

8

9

10

11

ABSTRACT

Statistical algorithms that are useful to aid in the identification of evolutionarily conserved amino acid positions within a family of proteins, and in the identification of interacting amino acid positions within a protein sequence are disclosed. The algorithms may also be useful in the analysis of other polymer sequences such as polysaccharides, lipids, deoxyribonucleic acid (DNA), and ribonucleic acid sequences (RNA), and, more specifically, in the analysis of DNA microarray data. Algorithms useful for analyzing the structural changes of perturbations to determine the mechanisms by which positions are coupled are also disclosed.